

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Coleman, Roger
Au-Young, Janice
Bandman, Olga
Seilhammer, Jeffrey

(ii) TITLE OF INVENTION: THROMBIN RECEPTOR HOMOLOG

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: INCYTE GENOMICS, INC.
(B) STREET: 3160 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: FastSeq for Windows Version 2.0

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/217,101
(B) FILING DATE: 21-DEC-1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/911,320
(B) FILING DATE: August 14, 1997

(viii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/467,125
(B) FILING DATE: 6-JUN-1995

(ix) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0041-3 DIV

(x) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-845-4170
(B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1143 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Liver
(B) CLONE: 86700

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAGAAGTC	ATACCATAAC	AATGACGACA	ACTTCAGTCA	GCAGCTGGCC	TTACTCCTCC	60
CACAGAATGC	GCTTTATAAC	CAATCATAGC	GACCAACCGC	CACAAAACCTT	CTCAGCAACA	120
CCAAATGTTA	CTACCTGTCC	CATGGATGAA	AAATTGCTAT	CTACTGTGTT	AACCACATCC	180
TACTCTGTTA	TTTTCATCGT	GGGACTGGTT	GGGAACATAA	TCGCCCCCTA	TGTATTTCTG	240
GGTATTCCAC	GTAAGAAGAA	TTCCATTCAA	ATTTATCTAC	TTAACGTAGC	CATTGCAGAC	300
CTCCTACTCA	TCTTCTGCCT	CCCTTTCCGA	ATAATGTATC	ATATTAACCA	AAACAAGTGG	360
ACACTAGGTG	TGATTCTGTG	CAAGGTTGTG	GGAACACTGC	TTTATATGAA	CATGTACATT	420
AGCATTATTT	TGCTTGGATT	CATCAGTTTG	GATCGCTATA	TAAAAATTAA	TCGGTCTATA	480
CAGCAACGGA	AGGCAATAAC	AACCAACAA	AGTATTTATG	TCTGTTGTAT	AGTATGGATG	540
CTTGCTCTTG	GTGGATTCCCT	AACTATGATT	ATTTTAAACAC	TTAAGAAAGG	AGGGCATAAT	600
TCCACAATGT	GTTTCCATTA	CAGAGATAAG	CATAACGCAA	AAGGAGAAGC	CATTTTTTAAC	660
TTCATTCTTG	TGGTAATGTT	CTGGCTAATT	TTCTTACTAA	TAATCCTTTC	ATATATTAAG	720
ATTGGGAAGA	ATCTATTGAG	GATTTCTAAA	AGGAGGTCAA	AATTTCCCTAA	TTCTGGTAAA	780
TATGCCACTA	CAGCTCGTAA	CTCCTTTATT	GTACTTATCA	TTTTTACTAT	ATGTGTGGGT	840
CCCTATCATG	CCTTTTCGATT	CATCTACATT	TCTTCACAGC	TAAATGTATC	ATCTTGCTAC	900
TGGAAAGAAA	TTGTTACAAA	AACCAATGAG	ATCATGCTGG	TTCTCTCATC	TTTCAATAGT	960
TGGTTAGATC	CAGTCATGTA	TTTCTGATG	TCCAGTAACA	TTGCGAAAAT	AATGTGCCAA	1020
CTTCTTTTTA	GACGATTTCA	AGGTGAACCA	AGTAGGAGTG	AAAGCACTTC	AGAATTTAAA	1080
CCAGGATACT	CCCTGCATGA	TACATCTGTG	GCAGGGAAAA	TACAGTCTAG	TTCTGAAAGT	1140
ACT						1143

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 381 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Arg	Ser	His	Thr	Ile	Thr	Met	Thr	Thr	Thr	Ser	Val	Ser	Ser	Trp
1				5					10					15	
Pro	Tyr	Ser	Ser	His	Arg	Met	Arg	Phe	Ile	Thr	Asn	His	Ser	Asp	Gln
			20					25					30		
Pro	Pro	Gln	Asn	Phe	Ser	Ala	Thr	Pro	Asn	Val	Thr	Thr	Cys	Pro	Met
		35					40					45			
Asp	Glu	Lys	Leu	Leu	Ser	Thr	Val	Leu	Thr	Thr	Ser	Tyr	Ser	Val	Ile
	50					55					60				
Phe	Ile	Val	Gly	Leu	Val	Gly	Asn	Ile	Ile	Ala	Pro	Tyr	Val	Phe	Leu

PF-0041-4 CON

65		70		75		80									
Gly	Ile	His	Arg	Lys	Arg	Asn	Ser	Ile	Gln	Ile	Tyr	Leu	Leu	Asn	Val
				85					90					95	
Ala	Ile	Ala	Asp	Leu	Leu	Leu	Ile	Phe	Cys	Leu	Pro	Phe	Arg	Ile	Met
			100					105					110		
Tyr	His	Ile	Asn	Gln	Asn	Lys	Trp	Thr	Leu	Gly	Val	Ile	Leu	Cys	Lys
		115				120						125			
Val	Val	Gly	Thr	Leu	Leu	Tyr	Met	Asn	Met	Tyr	Ile	Ser	Ile	Ile	Leu
	130					135					140				
Leu	Gly	Phe	Ile	Ser	Leu	Asp	Arg	Tyr	Ile	Lys	Ile	Asn	Arg	Ser	Ile
145					150					155					160
Gln	Gln	Arg	Lys	Ala	Ile	Thr	Thr	Lys	Gln	Ser	Ile	Tyr	Val	Cys	Cys
			165						170					175	
Ile	Val	Trp	Met	Leu	Ala	Leu	Gly	Gly	Phe	Leu	Thr	Met	Ile	Ile	Leu
		180					185					190			
Thr	Leu	Lys	Lys	Gly	Gly	His	Asn	Ser	Thr	Met	Cys	Phe	His	Tyr	Arg
		195				200						205			
Asp	Lys	His	Asn	Ala	Lys	Gly	Glu	Ala	Ile	Phe	Asn	Phe	Ile	Leu	Val
		210			215					220					
Val	Met	Phe	Trp	Leu	Ile	Phe	Leu	Leu	Ile	Ile	Leu	Ser	Tyr	Ile	Lys
225				230						235				240	
Ile	Gly	Lys	Asn	Leu	Leu	Arg	Ile	Ser	Lys	Arg	Arg	Ser	Lys	Phe	Pro
			245						250					255	
Asn	Ser	Gly	Lys	Tyr	Ala	Thr	Thr	Ala	Arg	Asn	Ser	Phe	Ile	Val	Leu
			260					265					270		
Ile	Ile	Phe	Thr	Ile	Cys	Val	Gly	Pro	Tyr	His	Ala	Phe	Arg	Phe	Ile
		275				280						285			
Tyr	Ile	Ser	Ser	Gln	Leu	Asn	Val	Ser	Ser	Cys	Tyr	Trp	Lys	Glu	Ile
	290				295					300					
Val	His	Lys	Thr	Asn	Glu	Ile	Met	Leu	Val	Leu	Ser	Ser	Phe	Asn	Ser
305				310						315				320	
Trp	Leu	Asp	Pro	Val	Met	Tyr	Phe	Leu	Met	Ser	Ser	Asn	Ile	Arg	Lys
			325						330					335	
Ile	Met	Cys	Gln	Leu	Leu	Phe	Arg	Arg	Phe	Gln	Gly	Glu	Pro	Ser	Arg
			340				345					350			
Ser	Glu	Ser	Thr	Ser	Glu	Phe	Lys	Pro	Gly	Tyr	Ser	Leu	His	Asp	Thr
		355				360						365			
Ser	Val	Ala	Gly	Lys	Ile	Gln	Ser	Ser	Ser	Glu	Ser	Thr			
	370					375					380				

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 425 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly Pro Arg Arg Leu Leu Leu Val Ala Ala Cys Phe Ser Leu Cys

1	5	10	15
Gly Pro Leu Leu Ser Ala Arg Thr Arg Ala Arg Arg Pro Glu Ser Lys			
20	25	30	
Ala Thr Asn Ala Thr Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro			
35	40	45	
Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp Glu Glu Lys Asn Glu Ser			
50	55	60	
Gly Leu Thr Glu Tyr Arg Leu Val Ser Ile Asn Lys Ser Ser Pro Leu			
65	70	75	80
Gln Lys Gln Leu Pro Ala Phe Ile Ser Glu Asp Ala Ser Gly Tyr Leu			
85	90	95	
Thr Ser Ser Trp Leu Thr Leu Phe Val Pro Ser Val Tyr Thr Gly Val			
100	105	110	
Phe Val Val Ser Leu Pro Leu Asn Ile Met Ala Ile Val Val Phe Ile			
115	120	125	
Leu Lys Met Lys Val Lys Lys Pro Ala Val Val Tyr Met Leu His Leu			
130	135	140	
Ala Thr Ala Asp Val Leu Phe Val Ser Val Leu Pro Phe Lys Ile Ser			
145	150	155	160
Tyr Tyr Phe Ser Gly Ser Asp Trp Gln Phe Gly Ser Glu Leu Cys Arg			
165	170	175	
Phe Val Thr Ala Ala Phe Tyr Cys Asn Met Tyr Ala Ser Ile Leu Leu			
180	185	190	
Met Thr Val Ile Ser Ile Asp Arg Phe Leu Ala Val Val Tyr Pro Met			
195	200	205	
Gln Ser Leu Ser Trp Arg Thr Leu Gly Arg Ala Ser Phe Thr Cys Leu			
210	215	220	
Ala Ile Trp Ala Leu Ala Ile Ala Gly Val Val Pro Leu Val Leu Lys			
225	230	235	240
Glu Gln Thr Ile Gln Val Pro Gly Leu Asn Ile Thr Thr Cys His Asp			
245	250	255	
Val Leu Asn Glu Thr Leu Leu Glu Gly Tyr Tyr Ala Tyr Tyr Phe Ser			
260	265	270	
Ala Phe Ser Ala Val Phe Phe Phe Val Pro Leu Ile Ile Ser Thr Val			
275	280	285	
Cys Tyr Val Ser Ile Ile Arg Cys Leu Ser Ser Ser Ala Val Ala Asn			
290	295	300	
Arg Ser Lys Lys Ser Arg Ala Leu Phe Leu Ser Ala Ala Val Phe Cys			
305	310	315	320
Ile Phe Ile Ile Cys Phe Gly Pro Thr Asn Val Leu Leu Ile Ala His			
325	330	335	
Tyr Ser Phe Leu Ser His Thr Ser Thr Thr Glu Ala Ala Tyr Phe Ala			
340	345	350	
Tyr Leu Leu Cys Val Cys Val Ser Ser Ile Ser Ser Cys Ile Asp Pro			
355	360	365	
Leu Ile Tyr Tyr Tyr Ala Ser Ser Glu Cys Gln Arg Tyr Val Tyr Ser			
370	375	380	
Ile Leu Cys Cys Lys Glu Ser Ser Asp Pro Ser Ser Tyr Asn Ser Ser			
385	390	395	400
Gly Gln Leu Met Ala Ser Lys Met Asp Thr Cys Ser Ser Asn Leu Asn			
405	410	415	
Asn Ser Ile Tyr Lys Lys Leu Leu Thr			
420	425		

PF-0041-4 CON

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

AAGGAGGGCATAATTCCACAATGTG

(25)

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

AAGGAGGGCATAATTCCACAATGTG

(25)